12

See 1473

EFFICIENT LINEAR SEARCH ALGORITHMS FOR THE LOGARITHMIC BARRIER FUNCTION

BY

WALTER MURRAY and MARGARET H. WRIGHT

TECHNICAL REPORT SOL 76-18
SEPTEMBER 1976

Systems Optimization Laboratory

Department of Operations Research

Stanford Univer[illegible]ity

Stanford California 94305

EFFICIENT LINEAR SEARCH ALGORITHMS FOR THE LOGARITHMIC BARRIER FUNCTION

by

Walter Murray* and Margaret H. Wright**

TECHNICAL REPORT SOL 76-18

September 1976

SYSTEMS OPTIMIZATION LABORATORY

DEPARTMENT OF OPERATIONS RESEARCH

Stanford University
Stanford, California

*Division of Numerical Analysis and Computing, National Physical Laboratory, Teddington, Middlesex, England.

**An earlier version of this paper was published as part of this author's Ph.D. dissertation.

Research and reproduction of this report were partially supported by the Office of Naval Research under Contract N00014-75-C-0865; the Energy Research and Development Administration E(04-3)-326 PA #18; and the National Science Foundation Grant DCR75-04544.

ABSTRACT

Linear search algorithms are developed for use when minimizing logarithmic barrier functions, whose one-dimensional behavior is in general modeled poorly by the low-order polynomial approximations of standard linear search procedures. The new methods are based on special approximating functions with a logarithmic singularity, and are designed to utilize the same information as procedures based on quadratic or cubic polynomials. Although the parameters of the special approximating functions depend nonlinearly on the available data, the determination of the parameters requires little additional work in comparison with polynomial fits. Use of the special approximating functions has led to a significant improvement in efficiency when minimizing logarithmic barrier functions, where efficiency is measured by the number of function (or function and gradient) evaluations required for termination of each linear search.

1. Introduction

An essential part of many algorithms for minimizing a function of several variables is the determination of a positive step to be taken from an initial point along a given search direction. At the new point, the values of the function and possibly its derivatives are required to satisfy some specified conditions with respect to these quantities at the original point. For example, in theoretical presentations of such algorithms, it is often stated that the step taken should minimize the function along the direction of search; the desired step in this case will be denoted by α^*, and is defined by:

$$\bar{F}(x + \alpha^* p) = \min_{\alpha} \bar{F}(x + \alpha p),$$

where x is the initial point, p is the search direction, and $\bar{F}(x)$ is the function to be minimized.

However, in most applications the conditions that define an acceptable step are less stringent, and will be satisfied by an approximation to α^* (see Osborne, 1972; Gill and Murray, 1974). An efficient method for obtaining such a step can be based on an iterative procedure to determine α^*, by generating iterates only until the required conditions hold. Thus, whatever the criteria to be satisfied, the underlying iterative process remains the same; what varies is the number of elements in the sequence that are actually computed.

The discussion to follow will be concerned only with the procedure by which successive estimates of the desired step length are generated, and not with the termination criteria. The terms "linear search" and "one-dimensional minimization" will be used to denote the iterative procedure common to all such algorithms, and do not imply that a close approximation to α^* is to be obtained.

The linear search procedure is of crucial importance to the success of most computational algorithms for unconstrained minimization; consequently, there has been considerable work on devising efficient and reliable methods for one-dimensional minimization, and on the careful implementation of such methods as computer routines (see Brent, 1973; Gill and Murray, 1974). These general-purpose linear search algorithms combine successive low-order polynomial interpolation or extrapolation with safeguards to ensure convergence and numerical stability. The performance of a linear search procedure is normally measured in terms of the number of function, or function and gradient evaluations required to locate a point that satisfies the prescribed termination criteria. When executing a linear search with respect to a particular nonlinear function, the efficiency of the algorithm will thus depend on the adequacy of the polynomial approximation to the given function. For minimizing general functions with no a priori information about the nonlinearities, linear search procedures based on polynomial approximation may be the most efficient, although inevitably the non-polynomial-like behavior of some functions will necessitate a large number of successive polynomial fits (and, hence, evaluations of the function).

This paper will be concerned with the design of linear search algorithms for a particular class of nonlinear functions--logarithmic barrier functions. A linear search with respect to these functions is required in several algorithms for constrained minimization, and yet their behavior is in general modeled poorly by low-order polynomials.

2. The Logarithmic Barrier Function

Barrier function methods were first introduced as a technique for transforming a nonlinearly constrained minimization problem into a sequence of unconstrained problems, where feasibility with respect to the problem constraints is maintained throughout (a full description is given in Fiacco and McCormick, 1968). The only barrier function to be considered here is the logarithmic barrier function, defined as follows: corresponding to the nonlinearly constrained minimization problem

$$\begin{aligned} &\text{minimize} \quad F(x), \qquad x \in E^n \\ &\text{subject to} \quad c_i(x) \geq 0, \quad i = 1, 2, \cdots, \ell, \end{aligned} \tag{2.1}$$

where $F(x)$ and $\{c_i(x)\}$ are prescribed nonlinear functions, the logarithmic barrier function is given by:

$$B(x, r) = F(x) - r \sum_{i=1}^{\ell} \ell n(c_i(x)), \tag{2.2}$$

where the positive scalar r is termed the "barrier parameter."

Interest in barrier function methods has waned in recent years, due to the development of more promising algorithms, many of which are based on extensions to the quadratic penalty function (see Fletcher, 1974; for a discussion of penalty functions, see Fiacco and McCormick, 1968). However, nearly all the more recent algorithms are non-feasible point methods, and hence are unsatisfactory for problems in which the objective and/or constraint functions are ill-defined or undefined outside the feasible region. In addition, non-feasible point methods may have other disadvantages. For example, the transformations employed may introduce spurious non-feasible solutions; furthermore, if such an algorithm is terminated prematurely, it often fails to provide a useable solution.

Many of the algorithmic developments based on the quadratic penalty function can be mirrored by utilizing the logarithmic barrier function. Two such suggestions have been made by Osborne (1972) and Murray and Wright (1976). For feasible point algorithms, the logarithmic barrier function, which has the virtue of maintaining feasibility, can often serve as a convenient "merit function" for a linear search, regardless of how the search direction is obtained. Therefore, it is worthwhile to consider how to exploit the special properties of the logarithmic barrier function in the design of a linear search algorithm.

The inadequacy of linear search procedures based on polynomial fits for minimizing barrier functions has been discussed by Fletcher and McCann (1969), Murray (1969), Lasdon et al. (1973), and Ryan (1974).

The defining characteristic of a barrier function is singularity at the boundary of the feasible region; since polynomials provide a poor approximation to a function with a singularity, the general approach has been to suggest alternative approximating functions with the same kind of singularity as the barrier function. The minimum of the special approximating function can be taken as an estimate of the minimum of the barrier function, to be used within the linear search procedure in exactly the same way as the estimated minima of polynomial approximations. However, it may not be possible to determine explicitly the coefficients and/or the minimum of the new approximating function, in contrast to these calculations for polynomials.

3. Fitting of Special Functions

3.1. Discussion

The special functions to be considered are designed to contain a single logarithmic singularity. This restriction to a particular barrier function contrasts with the general approach taken by Lasdon et al. (1973), where a special function is developed by the application of the form of the barrier function to linearized approximations to the objective function and the constraints. The approach taken here allows a much simpler solution to the problem, where the known form of the singularity is directly exploited. A single function is used to approximate the behavior of the barrier function along the search direction, and it is not necessary to make separate approximations to the objective function and the constraint functions.

The logarithmic barrier function can be evaluated only at feasible points, and is undefined beyond the first singularity along a particular direction. Therefore, only feasible points are considered in fitting the approximating functions, since the data to be used will be values of the barrier function and its gradient.

It will be assumed that we seek an approximation to the minimum along the search direction of

$$B(x, r) = F - r \sum_{i=1}^{\ell} \ell n(c_i),$$

where the barrier parameter, r, is known and fixed throughout the linear search. The special functions to be considered are the following, where the parameter θ represents the one-dimensional variation along the search direction:

(a) a linear function plus a logarithmic singularity, of the form:

$$f_L(\theta) = \hat{a} + \hat{b}\theta - r \,\ell n(\hat{d} - \theta);$$

(b) a quadratic function plus a logarithmic singularity, of the form:

$$f_Q(\theta) = \hat{a} + \hat{b}\theta + \hat{c}\theta^2 - r \,\ell n(\hat{d} - \theta).$$

Thses formulations attribute all the singular behavior to one location $(\theta = \hat{d})$; in fact, the singularity displayed by a barrier function depends in general upon the first constraint zero encountered along a given direction, and zeros that may occur beyond the first are irrelevant. This behavior becomes especially marked for small values of the barrier parameter, when the constraint functions bounded away from zero have almost no influence on the barrier function, and the nearest constraint has influence only very close to its zero along the given direction. A smooth function with a suitable damped singularity should, therefore, be an excellent model of a barrier function close to the boundary of the feasible region.

3.2. Linear Function Plus Logarithmic Singularity

The special function to be fitted is of the form:

$$f_L = \hat{a} + \hat{b}\theta - r\,\ell n(\hat{d} - \theta).$$

Differentiating f_L with respect to θ gives:

$$f_L' = \hat{b} + \frac{r}{\hat{d} - \theta},$$

and

$$f_L'' = \frac{r}{(\hat{d} - \theta)^2}.$$

Since the barrier function is decreasing at $\theta = 0$, this same condition will be required of f_L, so that

$$f_L'(0) = \hat{b} + \frac{r}{\hat{d}} < 0.$$

The singularity, $\hat{d}$, is assumed to be a positive step along the current direction, and, therefore, $\hat{b} < 0$.

A stationary point of f_L occurs at θ^* such that

$$f_L'(\theta^*) = 0,$$

or

$$\hat{b} + \frac{r}{\hat{d} - \theta^*} = 0$$

so that

$$\theta^* = \hat{d} + r/\hat{b} .$$

The expression for θ^* is equal to the value of $f_L'(0)$ times $(\hat{d}/\hat{b})$, and since $(\hat{d}/\hat{b})$ and $f_L'(0)$ are negative, θ^* must be positive. f_L'' is everywhere positive, so that the stationary point must be a minimum of f_L.

In order to specify the function f_L, three independent pieces of information are required about the behavior of the function to be fitted along the search direction; f_L will therefore be used in circumstances where a parabolic fit would normally be carried out to minimize a general function. Figure 1 illustrates a barrier function and its approximations by the function f_L and a parabola, both using the same data. It is clear that the special function gives a more accurate prediction than the parabola of the minimum of the barrier function.

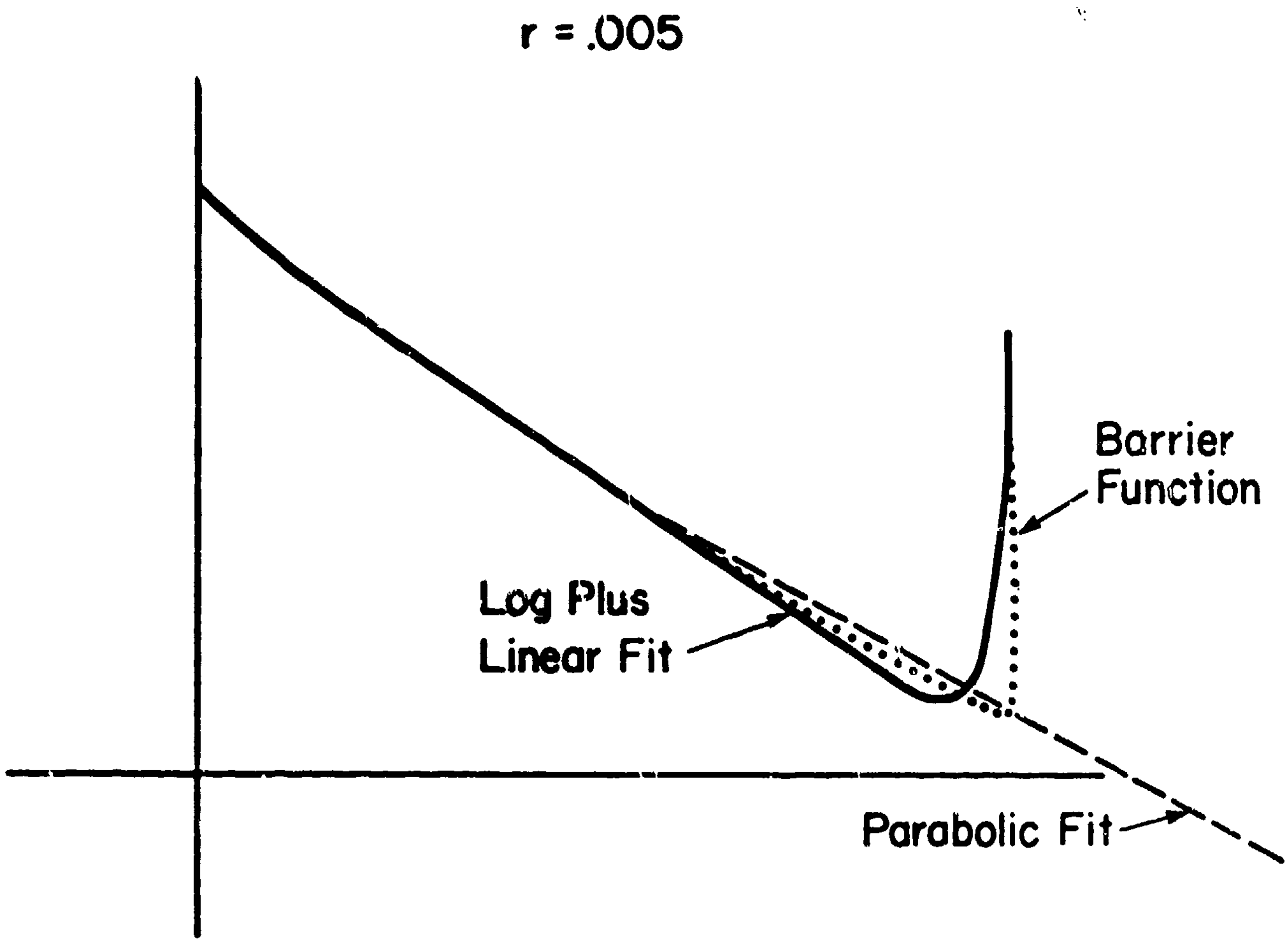

Figure 1

The special function f_L will be fitted with the same sets of data used by a typical parabolic line search, namely: (1) one function value, the corresponding projected gradient, and a second function value; (2) three function values. In the following discussion, f_i will denote the function value at θ_i, and g_i is the corresponding value of the projected gradient.

Case 1. Two function values, one gradient

The three unknown parameters of f_L -- $\hat{a}$, $\hat{b}$, and $\hat{d}$ -- must be solved for in terms of the known values. Assuming that $\theta_2 > \theta_1$, the equations specifying f_L are:

$$\text{(A)} \qquad f_1 = \hat{a} + \hat{b}\theta_1 - r\,\ell n(\hat{d} - \theta_1)$$

$$\text{(B)} \qquad g_1 = \hat{b} + \frac{r}{\hat{d} - \theta_1}$$

$$\text{(C)} \qquad f_2 = \hat{a} + \hat{b}\theta_2 - r\,\ell n(\hat{d} - \theta_2).$$

There is no loss of generality in assuming that $\theta_1 = 0$, so that from (A) and (B) we obtain:

$$\hat{a} = f_1 + r\,\ell n(\hat{d})$$

and

$$\hat{b} = g_1 - \frac{r}{\hat{d}}\,.$$

Substituting these values into (C) yields:

$$f_2 = f_1 + r\,\ell n(\hat{d}) + (g_1 - \frac{r}{\hat{d}})\theta_2 - r\,\ell n(\hat{d} - \theta_2),$$

or, after re-arranging:

$$(3.1) \qquad \ell n\left(\frac{\hat{d} - \theta_2}{\hat{d}}\right) + \frac{\theta_2}{\hat{d}} = \frac{1}{r}(\theta_2 g_1 - (f_2 - f_1)).$$

If a value $\hat{d}$, the location of the singularity, can be found which satisfies (3.1), the values $\hat{a}$, $\hat{b}$, and θ^* can then be computed from the previously derived relationships.

Consider (3.1) as a nonlinear equation in terms of the function

$$\Phi_1(d) = \ell n\left(\frac{d - \theta_2}{d}\right) + \frac{\theta_2}{d} + k_1,$$

where

$$k_1 = \frac{1}{r}(f_2 - f_1 - \theta_2 g_1).$$

The problem to be solved then becomes that of finding a solution $\hat{d}$ to satisfy

$$\Phi_1(d) = 0.$$

Because θ_2 is feasible, the value of the singularity, $\hat{d}$, must lie in the interval (θ_2, ∞). The function $\phi_1(d)$ has the following properties:

$$\phi_1 \to -\infty \quad \text{as} \quad d \to \theta_2 + ;$$

$$\phi_1 \to k_1 \quad \text{as} \quad d \to \infty ;$$

$$\phi_1' > 0 \quad \text{for all finite} \quad d;$$

$$\phi_1' \to \infty \quad \text{as} \quad d \to \theta_2 + ;$$

$$\phi_1' \to 0 \quad \text{as} \quad d \to \infty ;$$

and

$$\phi_1'' < 0 \quad \text{for all finite} \quad d.$$

These relations imply that there is a unique zero of ϕ_1 in (θ_2, ∞), provided that $k_1 > 0$, i.e., $(f_2 - f_1)/\theta_2 > g_1$. This requirement means that the function value at θ_2 must be larger than the linear approximation at θ_1 would have predicted, i.e., f_2 must lie in the shaded region of Figure 2.

Although the numerical solution of the equation $\phi_1(d) = 0$ would yield the desired value $\hat{d}$, ϕ_1 is an ill-behaved function, unsuitable for the usual zero-finding techniques such as Newton's method. Figure 3 illustrates the behavior of ϕ_1.

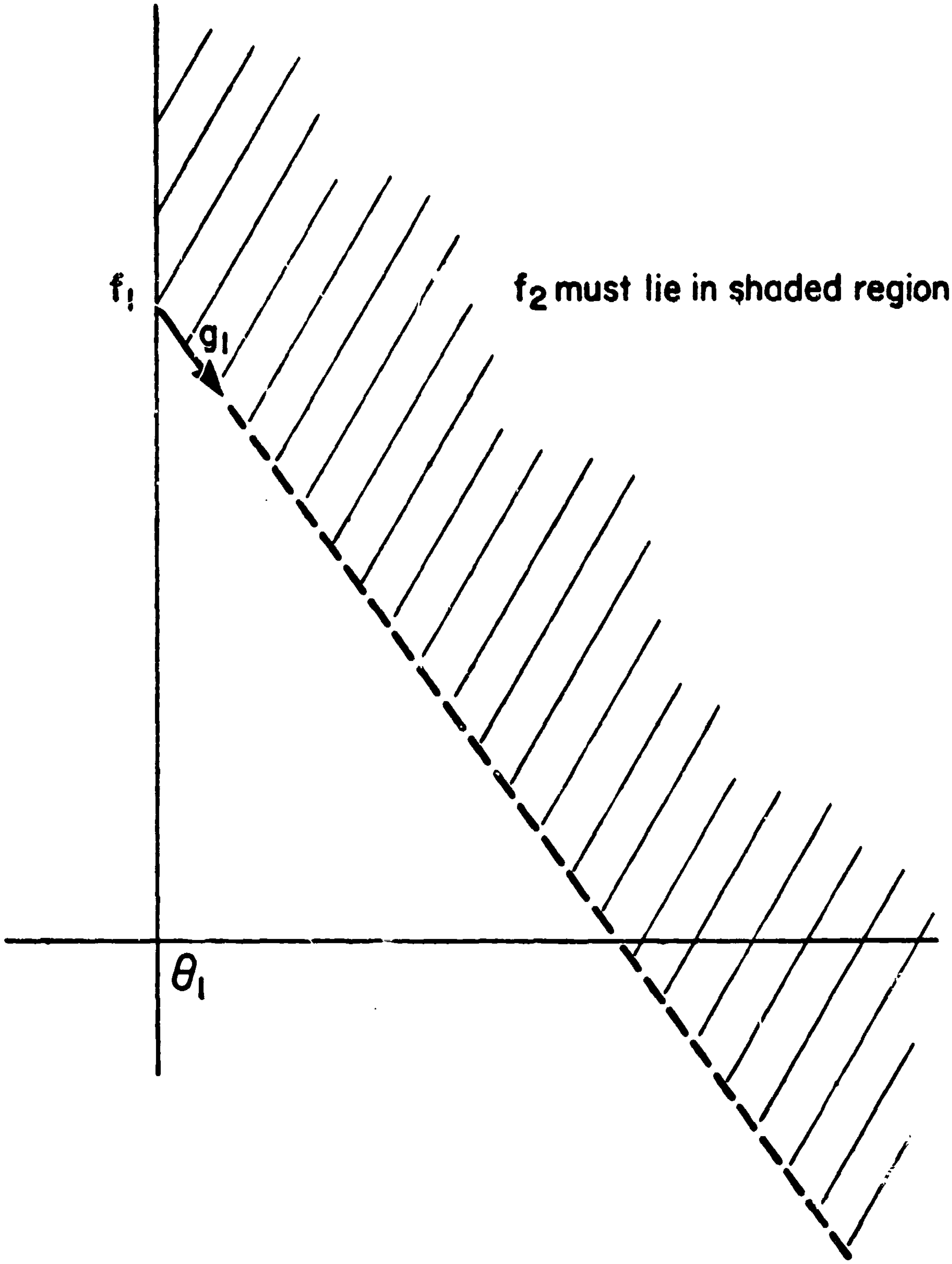

Figure 2

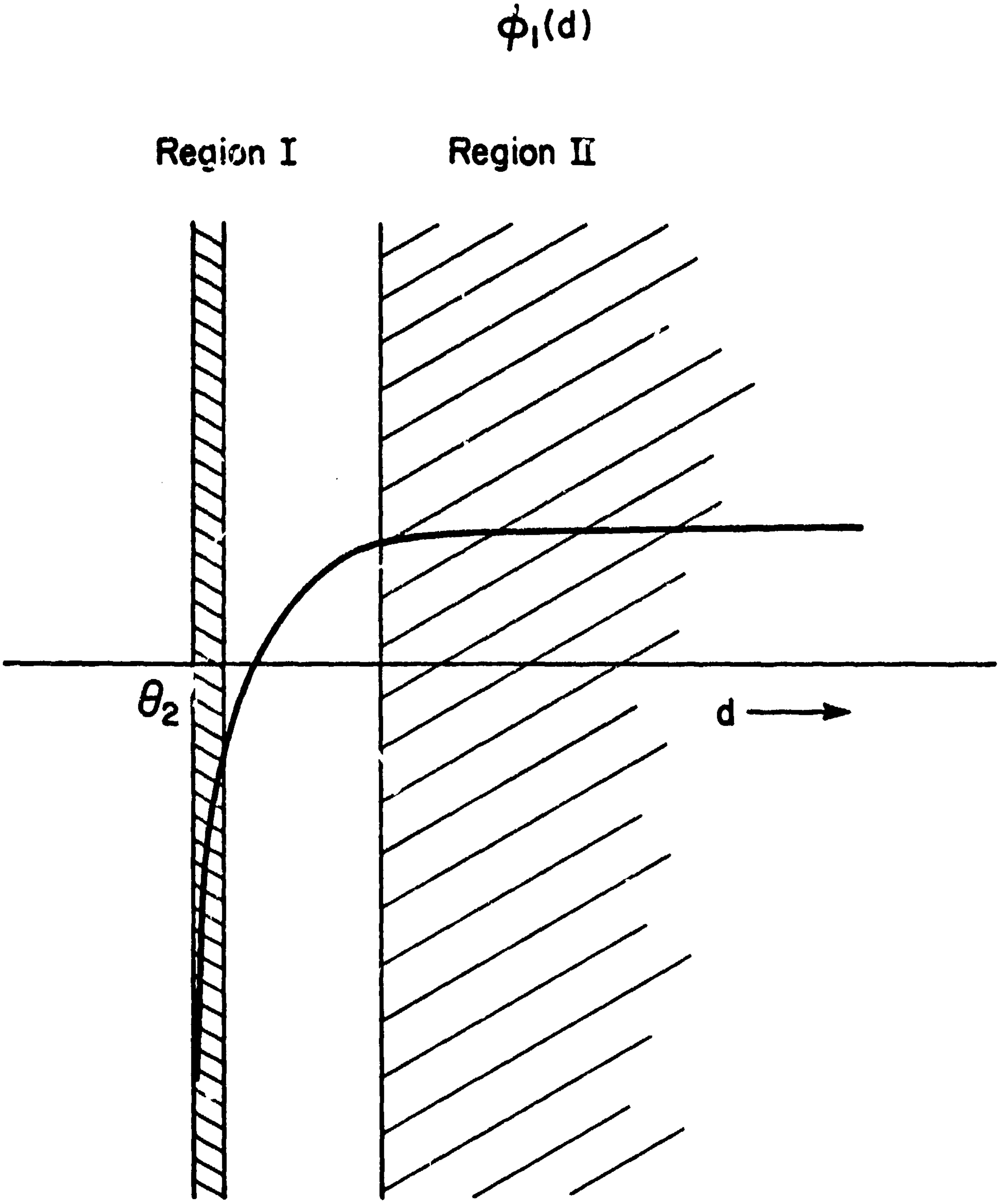

Figure 3

An attempt to apply Newton's method in Region I would cause very slow convergence to the zero, although all estimates would undershoot and hence could not diverge. In Region II, Newton's method might easily yield an estimate to the left of θ_2, and safeguards would need to be incorporated to prevent divergence.

To avoid these difficulties, the problem of solving $\Phi_1(d) = 0$ can be transformed into an equivalent problem that is easy to solve. If we define a new variable $y = \theta_2/d$, so that $0 < y < 1$ for the admissible range of d, the equation $\Phi_1(d) = 0$ can be written in terms of y as

$$\ln(1 - y) = -y - k_1 \ .$$

Taking exponentials of both sides gives the equation

$$1 - y = e^{-y} e^{-k_1} ,$$

and the value $\hat{y}$ which satisfies this equation is a zero of a new function,

$$\Psi_1(y) = 1 - y - e^{-y} e^{-k_1}.$$

Differentiating with respect to y gives:

$$\Psi_1' = -1 + e^{-y} e^{-k_1},$$

$$\Psi_1'' = -e^{-y} e^{-k_1},$$

and thus Ψ_1' and Ψ_1'' are negative for all y in the interval $(0, 1)$, if $k_1 > 0$. Furthermore,

$$\Psi_1(0) = 1 - e^{-k_1} > 0,$$

$$\Psi_1(1) = -e^{-1} e^{-k_1} < 0,$$

so that Ψ_1 has a unique zero $\hat{y}$ in $(0, 1)$ if $k_1 > 0$. Figures 4 and 5 illustrate the behavior of Ψ_1 for two values of k_1. Even for k_1 small, when the nonlinear portion of Ψ_1 is more significant, Ψ_1 is very close to linear in $(0, 1)$.

Because Ψ_1 is well-behaved, Newton's method will converge to $\hat{y}$ very rapidly. However, even further advantage can be taken of the form of Ψ_1 by noting that $\hat{y}$ satisfies the relation

$$k_1 = -\hat{y} - \ell n(1 - \hat{y}). \tag{3.2}$$

Since the right-hand side of (3.2) depends only on the parameter $\hat{y}$, and is independent of the problem data, the values of the function on the right-hand side can be tabulated for a set of y in $(0, 1)$. For a particular value of k_1, table lookup could then be used to determine a highly accurate estimate of $\hat{y}$. With a sufficiently large table, no iteration would be necessary to locate the solution to any desired accuracy; but since Newton's method converges rapidly,

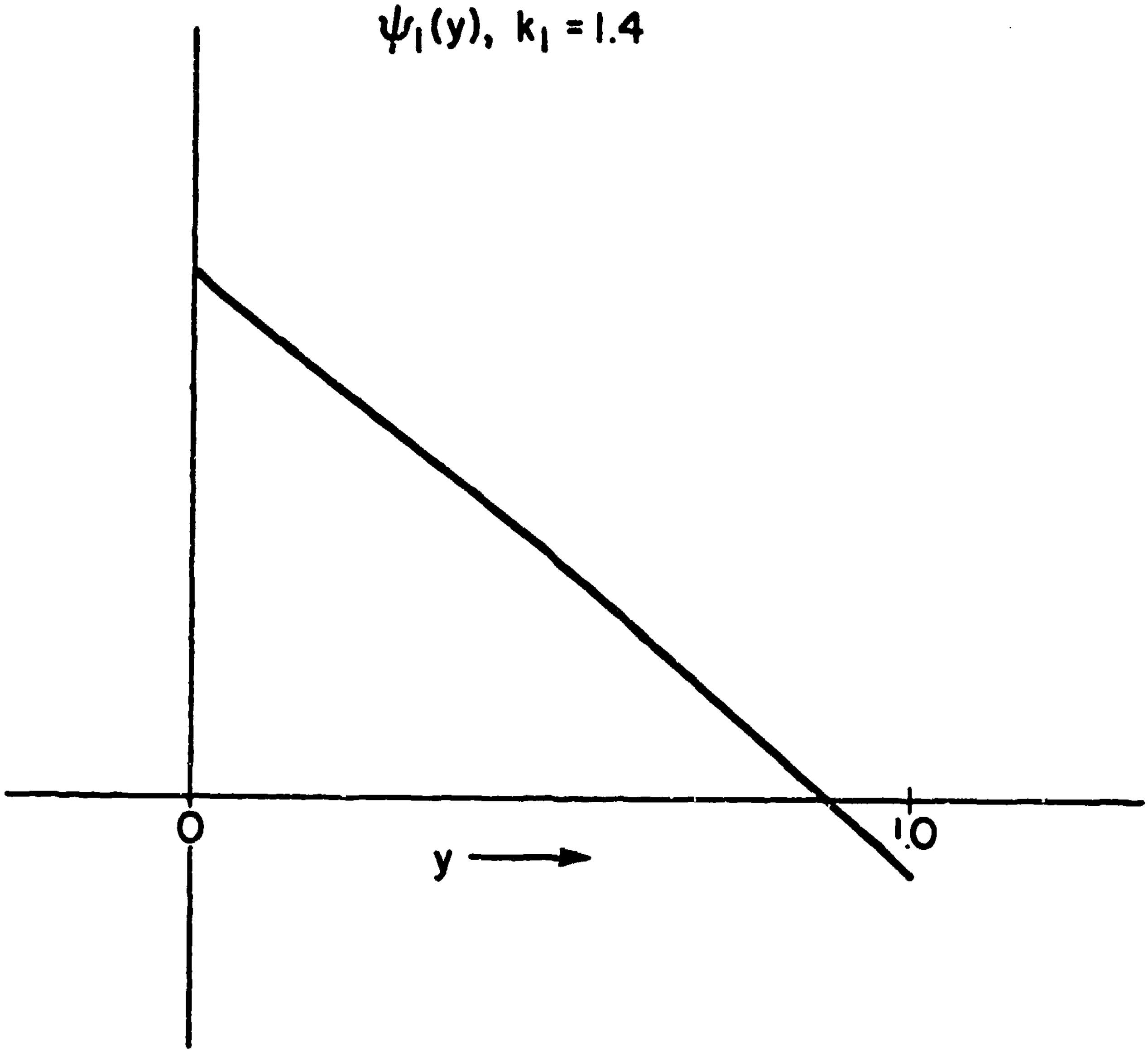

Figure 4

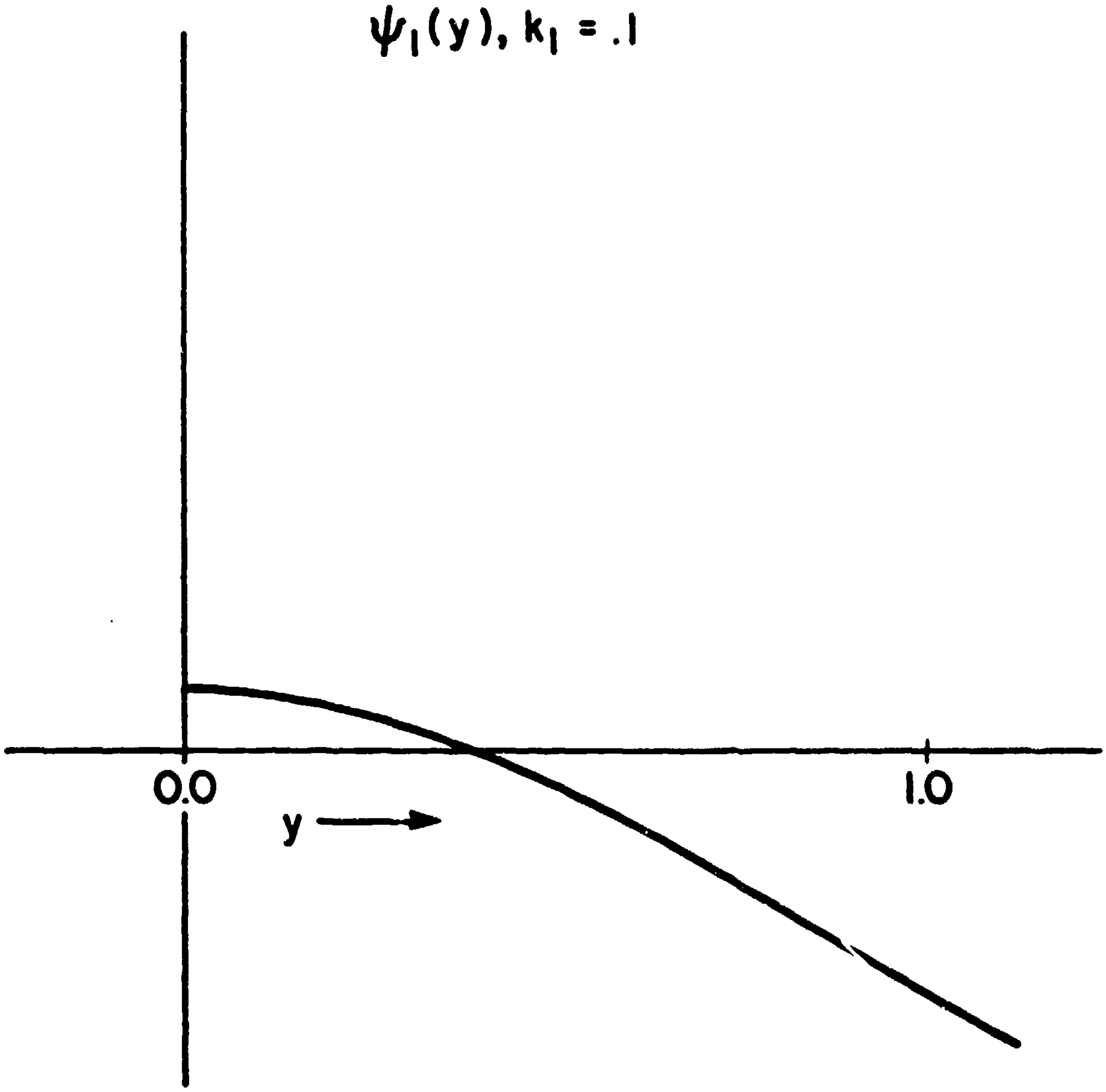

Figure 5

a small table plus a single iteration will locate $\hat{y}$ (and $\hat{d}$) to the required accuracy (e.g., seven values are sufficient for an accuracy of 10^{-5}).

Case 2. Three Function Values

Assuming that $\theta_3 > \theta_2 > \theta_1$, the three equations specifying the parameters $\hat{a}$, $\hat{b}$, and $\hat{d}$ of f_L are:

(A) $$f_1 = \hat{a} + \hat{b}\theta_1 - r\,\ell n(\hat{d} - \theta_1)$$

(B) $$f_2 = \hat{a} + \hat{b}\theta_2 - r\,\ell n(\hat{d} - \theta_2)$$

(C) $$f_3 = \hat{a} + \hat{b}\theta_3 - r\,\ell n(\hat{d} - \theta_3).$$

There is no loss of generality in assuming that $\theta_1 = 0$. Using equations (A) and (B) to eliminate $\hat{a}$, we obtain:

$$\frac{f_2 - f_1}{\theta_2} = \hat{b} - \frac{r}{\theta_2}\,\ell n\left(\frac{\hat{d} - \theta_2}{\hat{d}}\right)$$

and, similarly, using (A) and (C) gives:

$$\frac{f_3 - f_1}{\theta_3} = \hat{b} - \frac{r}{\theta_3}\,\ell n\left(\frac{\hat{d} - \theta_3}{\hat{d}}\right).$$

These two equations can then be manipulated to eliminate $\hat{b}$, yielding:

$$(3.3) \quad \frac{1}{\theta_3} \ln\left(\frac{\hat{d} - \theta_3}{\hat{d}}\right) - \frac{1}{\theta_2} \ln\left(\frac{\hat{d} - \theta_2}{\hat{d}}\right) + \frac{1}{r}\left(\frac{f_3 - f_1}{\theta_3} - \frac{f_2 - f_1}{\theta_2}\right) = 0,$$

a relationship which must be satisfied by $\hat{d}$, the location of the singularity.

Consider (3.3) as a nonlinear equation in terms of the function

$$\Phi_2(d) = \frac{1}{\theta_3} \ln\left(\frac{d - \theta_3}{d}\right) - \frac{1}{\theta_2} \ln\left(\frac{d - \theta_2}{d}\right) + k_2,$$

where

$$k_2 = \frac{1}{r}\left(\frac{f_3 - f_1}{\theta_3} - \frac{f_2 - f_1}{\theta_2}\right).$$

The problem to be solved then becomes that of finding a solution $\hat{d}$ to satisfy $\Phi_2(d) = 0$.

Because θ_2 and θ_3 are feasible, the value of $\hat{d}$ must lie in the interval (θ_3, ∞). The function Φ_2 has the following properties:

$$\begin{array}{ll} \Phi_2 \to -\infty & \text{as } d \to \theta_3 + ; \\ \Phi_2 \to k_2 & \text{as } d \to \infty ; \\ \Phi_2' > 0 & \text{for all finite } d; \\ \Phi_2' \to \infty & \text{as } d \to \theta_3 + ; \\ \Phi_2' \to 0 & \text{as } d \to \infty ; \end{array}$$

and

$$\Phi_2'' < 0 \qquad \text{for all finite } d.$$

These relations imply that there is a unique zero of Φ_2 in (θ_3, ∞) if $k_2 > 0$. This requirement has a similar interpretation to Case 1, i.e., the configuration of f_1, f_2, and f_3 must be as shown in Figure 6, so that the line joining f_3 and f_1 must lie above the line joining f_2 and f_1.

Because of its highly nonlinear behavior, Φ_2 is unsuited for application of the usual zero-finding techniques. The behavior of Φ_2 is illustrated in Figure 7.

To transform the problem of finding $\hat{d}$ to solve $\Phi_2(d) = 0$ into a computationally manageable form, we introduce the variables $\bar{d} = d/\theta_2$, so that $\bar{d} > 1$, and $\bar{\theta}_3 = \theta_3/\theta_2$, with $\bar{\theta}_3 > 1$; in essence, we consider θ_2 to be unity and scale θ_3 and d accordingly. If we multiply the equation $\Phi_2(d) = 0$ by θ_3 and substitute the variables $\bar{d}$ and $\bar{\theta}_3$, we obtain:

$$\ell n\left(1 - \frac{\bar{\theta}_3}{\bar{d}}\right) = \bar{\theta}_3 \; \ell n\left(1 - \frac{1}{\bar{d}}\right) - k_2\theta_3 \; . \tag{3.4}$$

Let the variable $y = \bar{\theta}_3/\bar{d}$, so that $0 < y < 1$ for all admissible d. Taking exponentials of both sides of (3.4) gives:

$$(1 - y) = \left(1 - \frac{y}{\bar{\theta}_3}\right)^{\bar{\theta}_3} e^{-k_2\theta_3} \; ,$$

and the value $\hat{y}$ which satisfies this relation is a zero of the function:

$$\psi_2(y) = 1 - y - \beta\left(1 - \frac{y}{\gamma}\right)^{\gamma} \; ,$$

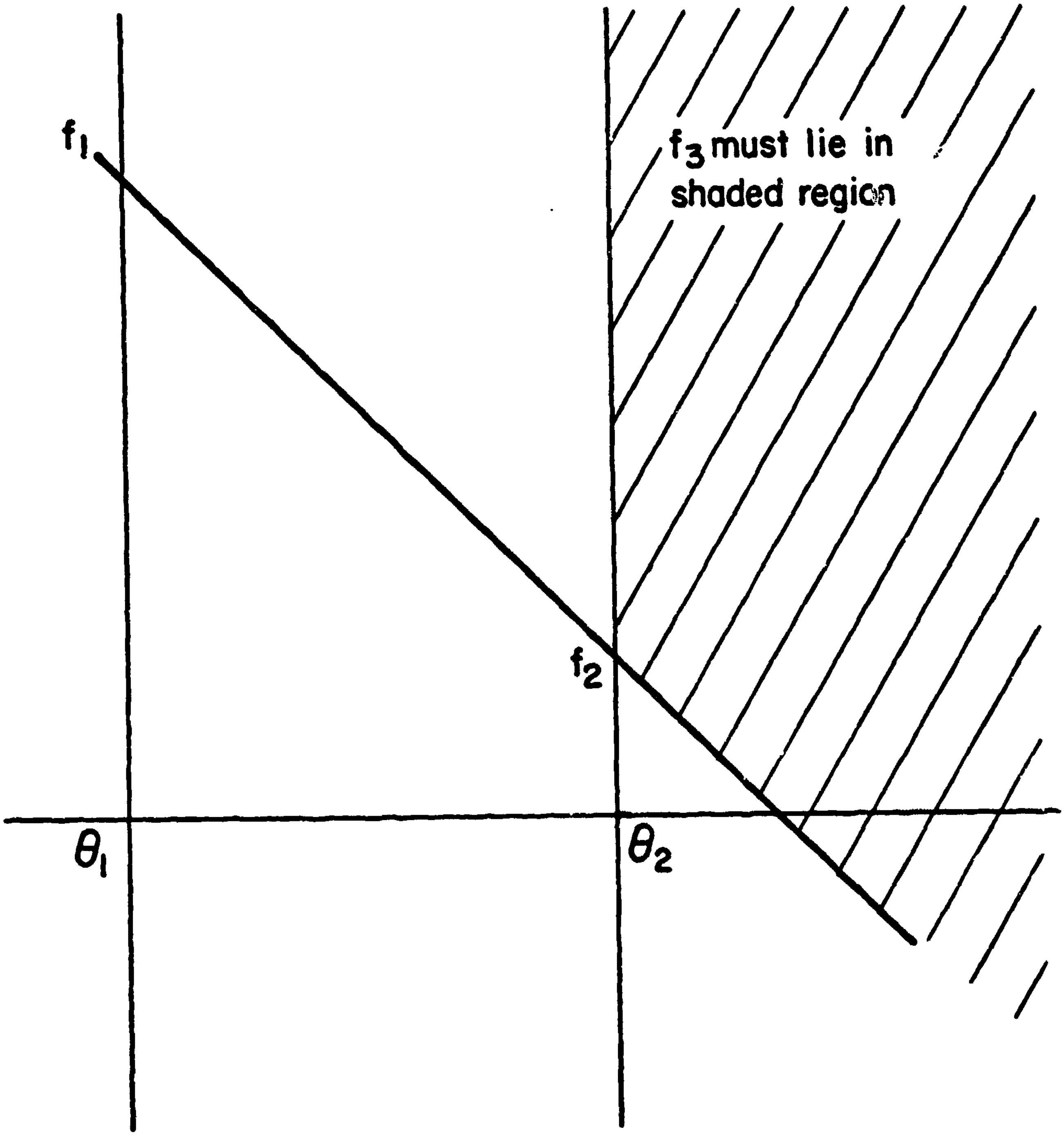

Figure 6

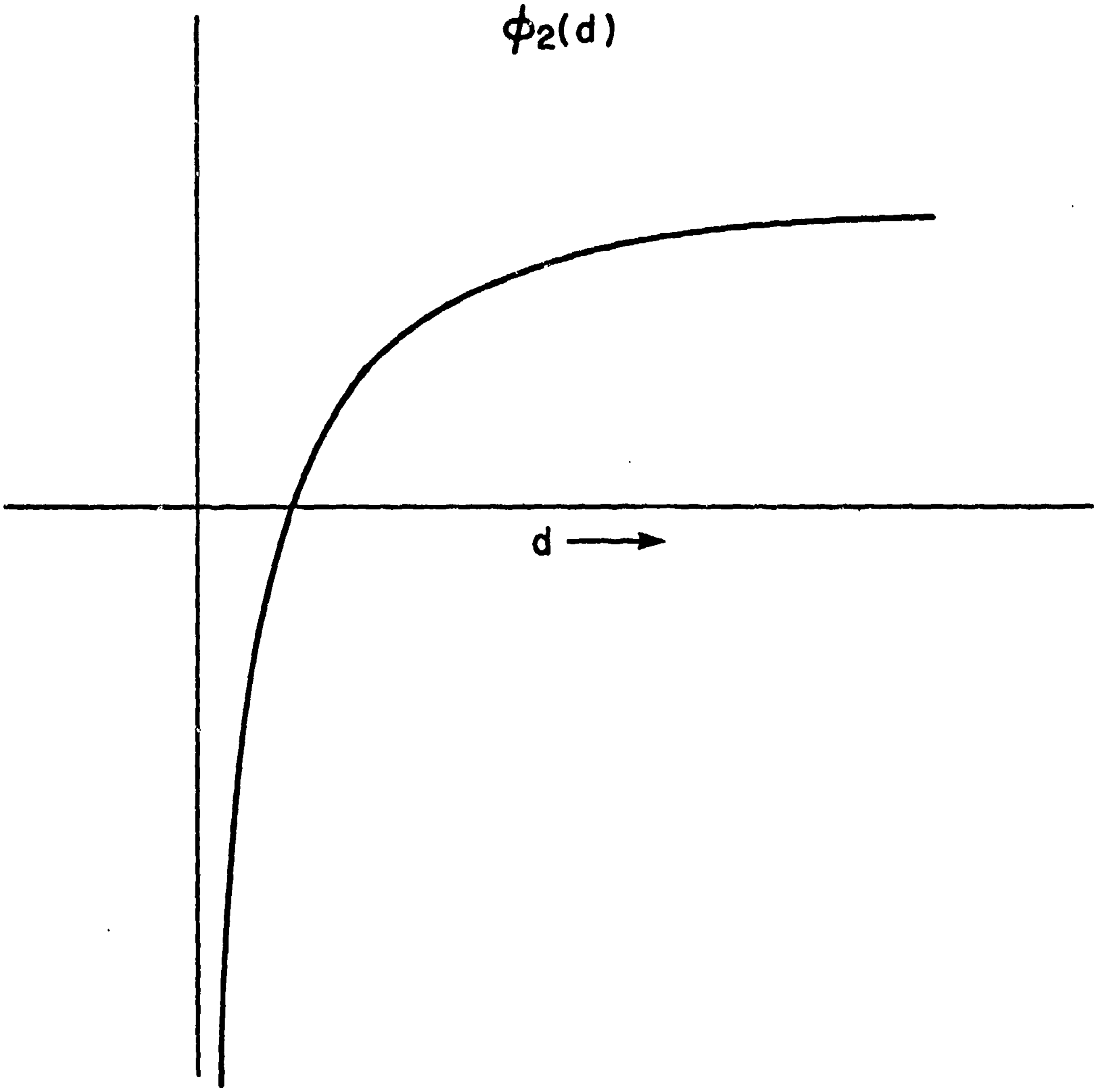

Figure 7

where

$$\gamma = \bar{\theta}_3 > 1, \quad \text{and} \quad \beta = e^{-k_2\theta_3} \qquad (\text{so that } 0 < \beta < 1 \text{ for } k_2 > 0).$$

Differentiating Ψ_2 with respect to y, we obtain

$$\Psi_2' = -1 + \beta\left(1 - \frac{y}{\gamma}\right)^{\gamma-1} ;$$

$$\Psi_2'' = -\beta \frac{\gamma - 1}{\gamma} \left(1 - \frac{y}{\gamma}\right)^{\gamma-2} .$$

Ψ_2 has the following properties:

$$\Psi_2(0) = 1 - \beta > 0 \qquad (y = 0 \text{ corresponds to } d \to \infty);$$

$$\Psi_2(1) = -\beta\left(1 - \frac{1}{\gamma}\right)^{\gamma} < 0 \qquad (y = 1 \text{ corresponds to } d \to \theta_3 +);$$

$$\Psi_2'(0) = -1 + \beta < 0$$

$$\Psi_2'(1) = -1 + \beta\left(1 - \frac{1}{\gamma}\right)^{\gamma-1} < 0,$$

and

$$\Psi_2'' < 0 \qquad \text{for } y \text{ in the interval } (0, 1).$$

These conditions imply that Ψ_2 has a unique zero in $(0, 1)$, which could be located by Newton's method. Ψ_2 is a well-behaved function, in contrast to Φ_2. The nonlinearity in Ψ_2 results from the expression $(1 - y/\gamma)^{\gamma}$, where $\gamma > 1$, and is the quotient θ_3/θ_2 from the original problem. The value of γ has been monitored during several runs with particular barrier functions; it never exceeded 2.0, and was usually in the range (1.01, 1.5). Even if γ were large, the expression $(1 - y/\gamma)^{\gamma}$ approaches e^{-y} as γ approaches ∞, and hence the function Ψ_2 takes on the form of Ψ_1 in Case 1. Furthermore, the value of $\beta (= e^{-k_2\theta_3})$ is usually small, so that the nonlinear component of Ψ_2 is even less significant.

Figures 8 and 9 illustrate the behavior of Ψ_2 for two sets of parameters (β, γ). Note that even for large β and γ, the nonlinearities do not have a significant effect.

It is not computationally convenient to obtain a tabulation of Ψ_2 in order to find an initial estimate of $\hat{y}$, since the influences of γ and y are not separable. However, Newton's method will converge rapidly from a reasonable starting point (for example, the estimate from Case 1).

3.3. Quadratic Function plus Logarithmic Singularity

The special function to be fitted is of the form:

$$f_0 = \hat{a} + \hat{b}\theta + \hat{c}\theta^2 - r\,\ell n(\hat{d} - \theta).$$

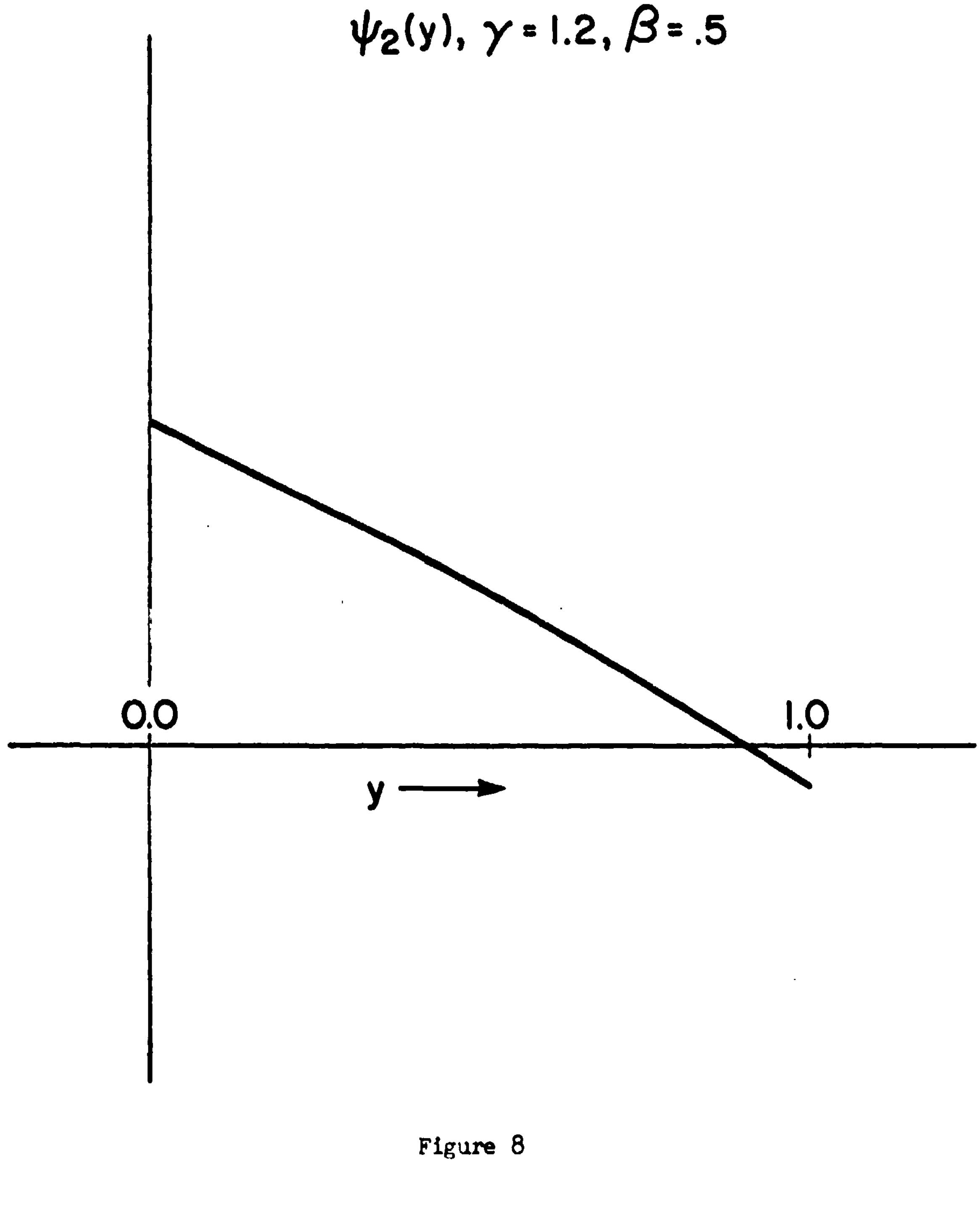

Figure 8

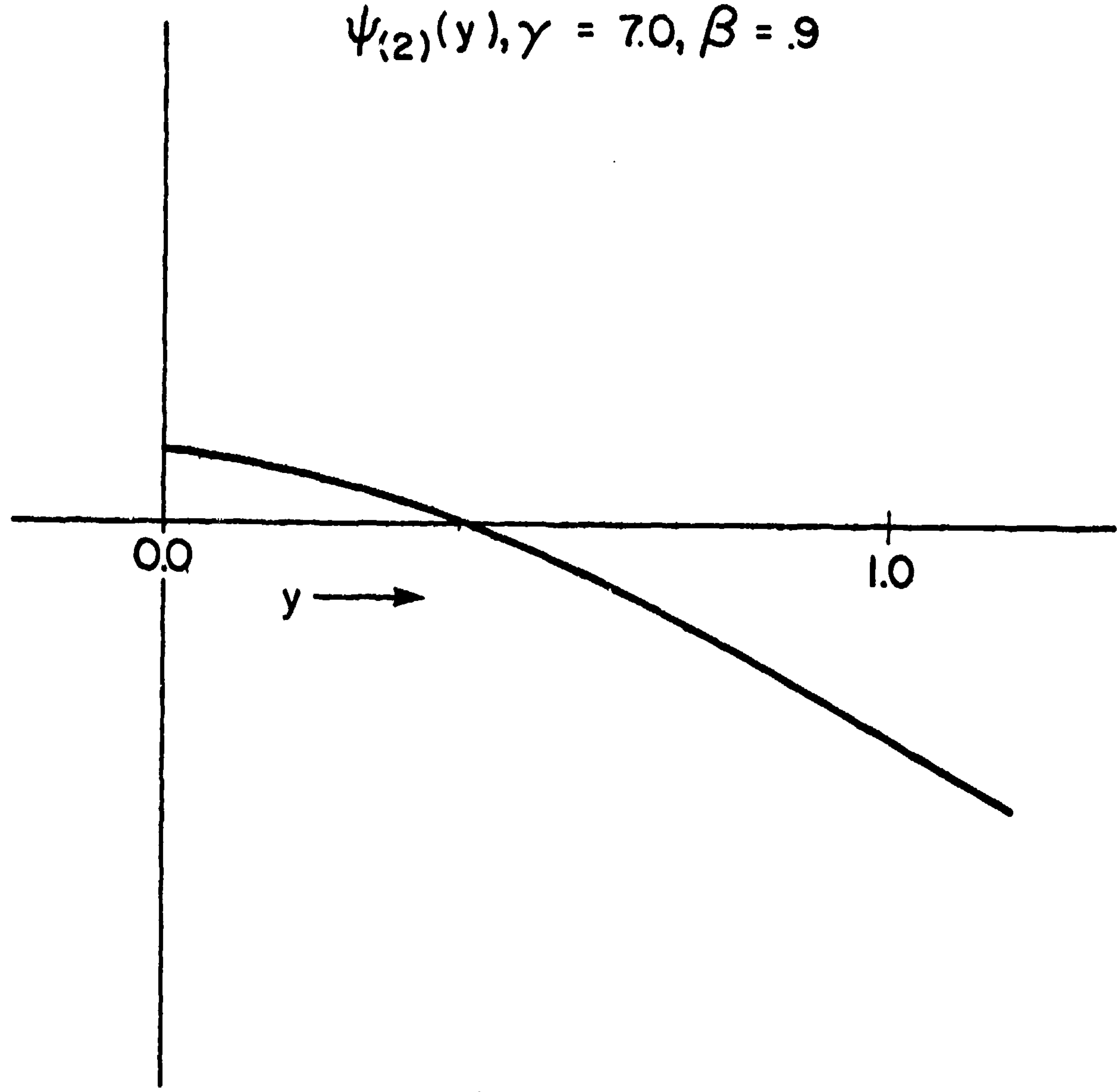

Figure 9

Differentiating f_Q with respect to θ gives:

$$f_Q' = \hat{b} + 2\hat{c}\theta + \frac{r}{\hat{d} - \theta},$$

$$f_Q'' = 2\hat{c} + \frac{r}{(\hat{d} - \theta)^2},$$

and

$$f_Q''' = \frac{2r}{(\hat{d} - \theta)^3}.$$

Since the barrier function is decreasing at $\theta = 0$, this same condition will be required of f_Q, so that $f_Q'(0) < 0$. Hence, there must exist a minimum of f_Q in the interval $(0, \hat{d})$, because $f_Q \rightarrow \infty$ as $\theta \rightarrow \hat{d}$.

A stationary point of f_Q will occur at θ^* where $f_Q'(\theta^*) = 0$, so that θ^* satisfies

$$\hat{b} + 2\hat{c}\theta^* + \frac{r}{\hat{d} - \theta^*} = 0.$$

This relation leads to a quadratic equation satisfied by θ^*,

$$2\hat{c}\theta^2 + (\hat{b} - 2\hat{c}\hat{d})\,\theta - r - \hat{b}\hat{d} = 0,$$

with two solutions which can be written in the alternative forms:

$$(3.5) \qquad \theta^* = \frac{2\hat{c}\hat{d} - \hat{b} \pm \sqrt{(\hat{b} - 2\hat{c}\hat{d})^2 + 8\hat{c}\,(r + \hat{b}\hat{d})}}{4\hat{c}} ,$$

or

$$(3.6) \qquad \theta^* = \frac{2\hat{c}\hat{d} - b \pm \sqrt{(\hat{b} + 2\hat{c}\hat{d})^2 + 8r\hat{c}}}{4\hat{c}} .$$

We now examine the properties of these two solutions in order to determine which will qualify as a suitable minimum, i.e., lie in the interval $(0, \hat{d})$. Since $f_Q'(0) < 0$ and $\hat{d} > 0$, the expression $\hat{b}\hat{d} + r$ is negative. Therefore, the sign of the term $8\hat{c}(r + \hat{b}\hat{d})$ is opposite to the sign of $\hat{c}$. Consider two cases:

<u>Case A:</u> $c > 0$ <u>(the quadratic term in f_Q has a positive coefficient)</u>.

The quantity $(\hat{b} + 2\hat{c}\hat{d})^2 + 8r\hat{c}$ involved in formula (3.6) for θ^* must be positive, so that two real roots exist. In formula (3.5) for θ^*, the quantity under the square root has magnitude less than $|2\hat{c}\hat{d} - \hat{b}|$, since $8\hat{c}(r + \hat{b}\hat{d})$ is negative. The expression $2\hat{c}\hat{d} - \hat{b}$ is positive since $\hat{c} > 0$, $\hat{d} > 0$, $\hat{b} < 0$, and, therefore, bot roots are positive. From formula (3.6) for θ^*, we note that the quantity under the square root exceeds $|\hat{b} + 2\hat{c}\hat{d}|$ in magnitude.

Case A-i: $\hat{b} + 2\hat{c}\hat{d} > 0$.

The value of θ^* corresponding to the positive square root satisfies:

$$\theta^*_+ > \frac{(2\hat{c}\hat{d} - \hat{b} + \hat{b} + 2\hat{c}\hat{d})}{4\hat{c}} = \hat{d},$$

and hence is unacceptable.

The value of θ^* for the negative square root satisfies:

$$\theta^*_- < \frac{2\hat{c}\hat{d} - \hat{b} - (\hat{b} + 2\hat{c}\hat{d})}{4\hat{c}} = -\frac{\hat{b}}{2\hat{c}}$$

Since $\hat{b} + 2\hat{c}\hat{d}$ is positive, and $\hat{c}$ is positive, it follows that $\hat{d} > -(\hat{b}/2\hat{c})$, and θ^*_- is an acceptable choice for a minimum.

Case A-ii: $\hat{b} + 2\hat{c}\hat{d} < 0$ (i.e., $|\hat{b} + 2\hat{c}\hat{d}| = -\hat{b} - 2\hat{c}\hat{d}$).

Here, the root corresponding to the positive square root satisfies:

$$\theta^*_+ > \frac{(2\hat{c}\hat{d} - \hat{b} - \hat{b} - 2\hat{c}\hat{d})}{4\hat{c}} = -\frac{\hat{b}}{2\hat{c}}.$$

Since $\hat{b} + 2\hat{c}\hat{d} < 0$, $\hat{d}$ must be less than $-(\hat{b}/2\hat{c})$; thus, $\theta^*_+ > \hat{d}$, and is unacceptable.

The solution corresponding to the negative square root satisfies:

$$\theta^*_- < \frac{(2\hat{c}\hat{d} - \hat{b} + (\hat{b} + 2\hat{c}\hat{d})}{4\hat{c}} = \hat{d},$$

and is acceptable.

Case B: $\hat{c} < 0$ (the coefficient of the quadratic term in f_Q is negative).

The quantity under the square root must be positive since $\hat{c} < 0$, $r + \hat{b}\hat{d} < 0$, and thus there are two real roots. In formula (3.5) for θ^*, the quantity under the square root exceeds $2\hat{c}\hat{d} - \hat{b}$ in magnitude, and hence there must be one positive and one negative root. Because of the sign reversal caused by division by $\hat{c}$, the positive square root corresponds to a value $\theta^* < 0$, and can be eliminated from consideration. The only positive θ^* can be written as:

$$\theta^* = \frac{\hat{b} - 2\hat{c}\hat{d} + \sqrt{(\hat{b} + 2\hat{c}\hat{d})^2 + 8r\hat{c}}}{4|\hat{c}|}.$$

The quantity under the square root has magnitude less than $|\hat{b} + 2\hat{c}\hat{d}|$, so that θ^* satisfies:

$$\theta^* < \frac{\hat{b} - 2\hat{c}\hat{d} - \hat{b} - 2\hat{c}\hat{d}}{4|\hat{c}|} = -\frac{4\hat{c}\hat{d}}{4|\hat{c}|} = \hat{d}.$$

For Cases A and B, then, assuming that $f_Q'(0) < 0$, the solution θ^* corresponding to the negative square root in formula (3.5) or (3.6) will be taken as the minimum of f_Q.

In order to specify the function f_Q, four independent pieces of information are required about the behavior of the function to be fitted along the search direction; f_Q will therefore be used in circumstances where a cubic fit would normally be carried out to minimize a general function. Figure 10 illustrates a barrier function and its approximations by the function f_Q and a cubic, both using the same data. The more accurate modeling of the barrier function by the special function is quite noticeable.

The special function f_Q will be fitted with the same set of data used by a typical cubic line search, namely, two function values and the corresponding two gradients. As in the previous discussion, f_i denotes the function value at θ_i, and g_i is the corresponding value of the projected gradient.

The four unknown parameters of f_Q - $\hat{a}$, $\hat{b}$, $\hat{c}$, and $\hat{d}$ -- must be solved for in terms of the known values. Assuming that $\theta_2 > \theta_1$, the equations specifying f_Q are:

(A) $$f_1 = \hat{a} + \hat{b}\theta_1 + \hat{c}\theta_1^2 - r\,\ln(\hat{d} - \theta_1)$$

(B) $$g_1 = \hat{b} + 2\hat{c}\theta_1 + \frac{r}{(\hat{d} - \theta_1)}$$

(C) $$f_2 = \hat{a} + \hat{b}\theta_2 + \hat{c}\theta_2^2 - r\,\ln(\hat{d} - \theta_2)$$

(D) $$g_2 = \hat{b} + 2\hat{c}\theta_2 + \frac{r}{(\hat{d} - \theta_2)} .$$

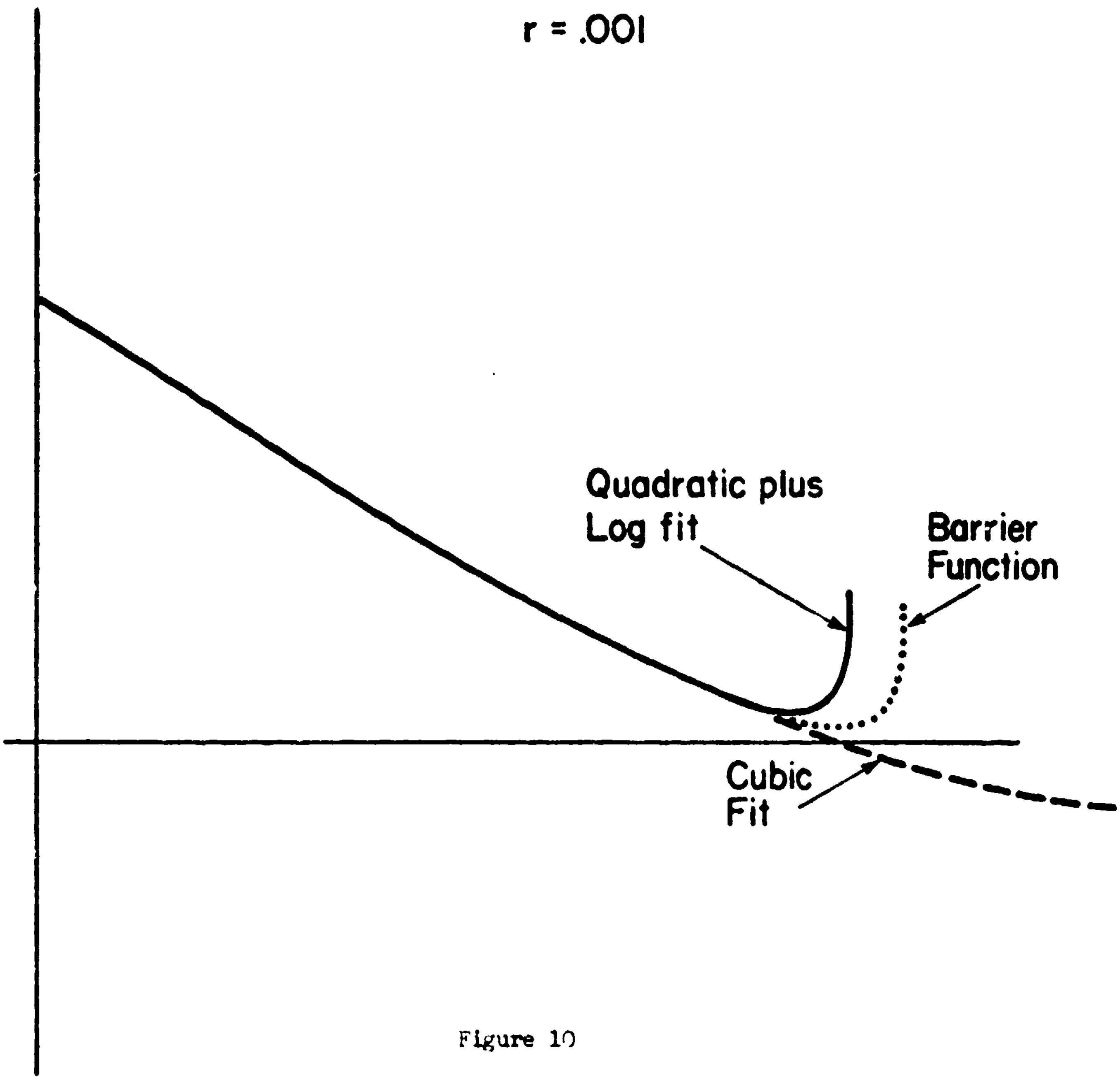

Figure 10

There is no loss of generality in assuming that $\theta_1 = 0$, so that from (A) and (B) we obtain expressions for the coefficients $\hat{a}$ and $\hat{b}$ in terms of $\hat{d}$:

$$\hat{a} = f_1 + r\, \ell n(\hat{d})$$

$$\hat{b} = g_1 - \frac{r}{\hat{d}} .$$

Substituting for $\hat{b}$ in equation (D), we obtain an expression for the parameter $\hat{c}$ in terms of $\hat{d}$:

$$\hat{c} = \frac{1}{2\theta_2} \left(g_2 - g_1 + \frac{r}{\hat{d}} - \frac{r}{\hat{d} - \theta_2}\right) .$$

These three expressions for $\hat{a}$, $\hat{b}$, and $\hat{c}$ can then be substituted into equation (C), to obtain the following:

$$(3.7) \qquad \ell n(\hat{d} - \theta_2) - \ell n(\hat{d}) + \frac{\theta_2}{2} \left(\frac{1}{\hat{d}} + \frac{1}{\hat{d} - \theta_2}\right)$$

$$= \frac{1}{r} \left(\frac{\theta_2}{2} (g_1 + g_2) - (f_2 - f_1)\right) .$$

If a value $\hat{d}$, the location of the singularity, can be found which satisfies (3.7), the values $\hat{a}$, $\hat{b}$, $\hat{c}$, and θ^* can then be computed from the previously derived relationships.

Consider (3.7) as a nonlinear equation in terms of the function

$$\Phi_3(d) = \ell n\left(\frac{d - \theta_2}{d}\right) + \frac{\theta_2}{2}\left(\frac{1}{d} + \frac{1}{d - \theta_2}\right) - k_3,$$

where

$$k_3 = \frac{1}{r}\left(\frac{\theta_2}{2}(g_1 + g_2) - (f_2 - f_1)\right).$$

The problem to be solved then becomes that of finding a solution $\hat{d}$ to satisfy

$$\Phi_3(d) = 0. \tag{3.8}$$

If we introduce the variable $z = 1 - \theta_2/d = (d - \theta_2)/d$, where $0 < z < 1$, we can then write the equation $\Phi_3(d) = 0$ in terms of the variable z as:

$$\ell n(z) + \frac{1}{2}\left(\frac{1}{z} - z\right) = k_3 . \tag{3.9}$$

The nonlinear equation (3.9) could be solved for a suitable $\hat{z}$, but the function represented is extremely ill-behaved. As $z \to 0$, the logarithm term is approaching $(-\infty)$, while the reciprocal term is simultaneously approaching $(+\infty)$; it is evident that (3.9) is quite unsuitable for purposes of computation.

However, the relationship (3.9) can be transformed into an equivalent form that is computationally reasonable. A further change of variable is made:

$$v = \ell n(z) ,$$

so that $z = e^{v}$, and $\frac{1}{z} = e^{-v}$; note that v will be nonpositive since $0 < z < 1$.

When written in terms of v, the relation (3.9) becomes:

$$v + \frac{1}{2}(e^{-v} - e^{v}) = k_3.$$

Since

$$\sinh(v) = \frac{1}{2}(e^{v} - e^{-v}),$$

the final result is:

$$v - \sinh(v) = k_3.$$

The value $\hat{v}$ that satisfies this relation is a zero of the function Ψ_3, where

$$\Psi_3(v) = k_3 + \sinh(v) - v.$$

Differentiating Ψ_3 with respect to v, we obtain:

$$\Psi_3' = \cosh(v) - 1$$

$$\Psi_3'' = \sinh(v),$$

so that the function Ψ_3 has the following properties:

$$\Psi_3(0) = k_3 \qquad (v = 0 \text{ corresponds to } d \to \infty);$$

$$\lim_{v \to -\infty} \Psi_3(v) = -\infty;$$

$$\Psi_3' > 0 \qquad \text{for } v < 0;$$

$$\Psi_3'(0) = 0;$$

$$\lim_{v \to -\infty} \Psi_3'(v) = \infty;$$

$$\Psi_3'' < 0 \qquad \text{for } v < 0.$$

These conditions imply that Ψ_3 has a unique zero in $(-\infty, 0)$ if $k_3 > 0$. This requirement means that the average of the gradients at θ_1 and θ_2 must exceed the slope of the straight line joining f_1 and f_2. If the function to be approximated were quadratic, the average of the slopes at θ_1 and θ_2 would exactly equal the slope of the line joining f_1 and f_2. The condition $k_3 > 0$ thus implies that the function to be approximated is rising more rapidly than a quadratic.

Figure 11 illustrates the behavior of Ψ_3.

Although the function Ψ_3 is unbounded below as $v \to -\infty$ (i.e., when $z \to 0$, or $d \to \theta_2 +$), this property does not cause any computational difficulties in the current context. The unbounded behavior of Ψ_3 occurs when the estimated value of the singularity is very close to θ_2; if a tolerance, say ϵ, is specified such that any estimate of $\hat{d}$ is required to satisfy $\hat{d} \geq \theta_2/(1 - \epsilon)$, then the variable z is bounded below by ϵ, and the variable v is bounded below by $-M$, $M > 0$, where $M = -\ell n(\epsilon)$. If values of v are restricted to the range $(-M, 0)$, the region where Ψ_3 is unbounded is eliminated. If k_3 is very large, it is possible that the value $\Psi_3(-M)$ will not be negative for the particular value of M chosen, and hence no zero of Ψ_3 will exist in $(-M, 0)$. Under these circumstances, we simply accept $\hat{v} = -M$ as the solution, so that $\hat{d} = \theta_2/(1 - \epsilon)$.

We can easily solve the equation $\Psi_3(v) = 0$ with Newton's method, considering the following properties of Ψ_3. Since $\Psi_3'' < 0$ throughout the interval of interest, if the starting point is chosen so that $\Psi_3 < 0$, the Newton iterates will undershoot the solution, and cannot diverge. Furthermore, the condition $\Psi_3(v) = 0$ can be written as:

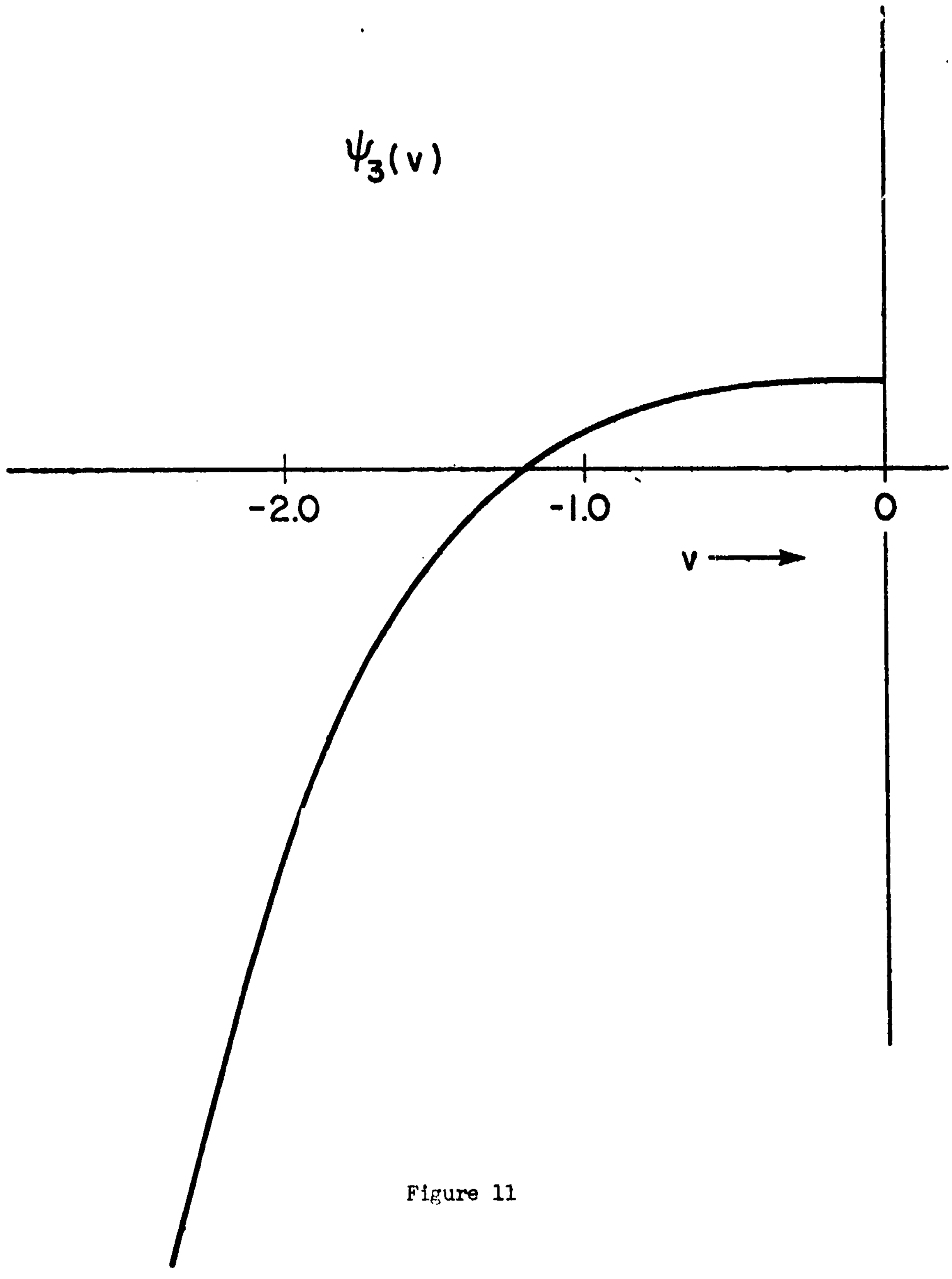

Figure 11

$$k_3 = v - \sinh(v),$$

and the expression on the right-hand side is independent of the problem data. Hence, the function $v - \sinh(v)$ can be tabulated for v in the range $(-M, 0)$, and by table lookup using the value k_3, a highly accurate initial estimate of v, with $\Psi_3 < 0$, can be obtained.

The Newton procedure converges extremely rapidly except when the value of v is close to zero, because $\Psi_3'(0) = 0$. This situation is quite unlikely because the estimated singularity would then be much larger than θ_2. For completeness, however, we note that the problem can be solved successfully even for very small v. The function $\Psi_3(v)$ can be written:

$$\begin{aligned}
\Psi_3(v) &= k_3 + \sinh(v) - v \\
&= k_3 + \frac{1}{2}(e^v - e^{-v}) - v \\
&= k_3 + \frac{1}{2}\left(1 + v + \frac{v^2}{2} + \frac{v^3}{6} + \cdots - \left(1 - v + \frac{v^2}{2} - \frac{v^3}{6} + \cdots\right)\right) - v \\
&= k_3 + \frac{1}{2}\left(2v + \frac{2v^3}{6} + \frac{2v^5}{120} + \cdots\right) - v \\
&= k_3 + \frac{v^3}{6} + O(v^5).
\end{aligned}$$

For small v, the equation $\Psi_3(v) = 0$ thus essentially becomes the condition

$$k_3 + \frac{v^3}{6} = 0,$$

with explicit solution $v^* = (-6k_3)^{1/3}$. Because we are ignoring negative higher-order terms, this value v^* will be to the left of the correct $\hat{v}$, and Newton's method cannot diverge. However, the estimate v^* is so accurate that no iteration at all is necessary to obtain an acceptable solution.

4. Implementation

The safeguarded linear searches based on quadratic or cubic interpolation (cf. Gill and Murray, 1974) have been modified for use with the logarithmic barrier function by allowing interpolation with the special functions described. Several rather complicated modifications are required in order to create an efficient algorithm. If no constraint is decreasing along the current search direction, or if no constraint approaches zero until sufficiently far beyond the starting point, then the singularity introduced to preserve feasibility will have no significant effect on the location of the minimum, and the usual linear search procedure should be followed. There is no computationally reasonable way to determine *a priori* whether these conditions exist because the constraints and objective function may be highly nonlinear, and the effort expended to compare the location of the nearest constraint zero with the prediction of the barrier function's minimum might be better used directly to minimize the barrier function. The procedure to be described seems to be a satisfactory compromise between excessive safeguards and unwarranted assumptions of linearity or smoothness.

4.1. Initial Step

The choice of the first step along the search direction at which the function is to be evaluated is affected by the possibility that a constraint may become nonpositive if the usual choice of step for the algorithm is taken. For example, with a Newton-type method, the initial step taken along the search direction is unity; for a quasi-Newton method, there is normally a procedure associated with the method for choosing the initial step. Let α_u denote the initial step that would be taken for a particular unconstrained algorithm if used to minimize the barrier function along the given direction. If a constraint might become zero at $\hat{\alpha} < \alpha_u$, clearly a shorter step than α_u should be taken. One possible method for determining the initial step is to find a highly accurate estimate of the step to the nearest zero of a constraint, say $\hat{\alpha}$, and test whether $\hat{\alpha} < \alpha_u$. A subroutine is available that will, with high reliability, locate the zero of the nearest constraint by use of a combination of safeguarded zero-finding techniques. However, locating $\hat{\alpha}$ generally requires several constraint evaluations, and it may turn out that $\hat{\alpha}$ exceeds α_u or is very close to α_u, so that these evaluations were essentially redundant. One might think of using the zero-finding technique until the zero has been shown conclusively to lie beyond α_u, but this approach involves quite complicated housekeeping, and, more significantly, may still require constraint calculations that do not advance the computation.

With the "compromise" algorithm, the initial step $\alpha^{(0)}$ to be taken along the search direction, p, is computed as follows:

1. Compute α_u, the step normally taken by the unconstrained method;
2. Compute the gradient of each constraint along p, i.e., $a_i^T p$, where a_i is the gradient of c_i. For all i such that this gradient is negative, i.e., the i-th constraint is locally decreasing along p, compute $\alpha_i = -c_i/a_i^T p$, the predicted Newton step to the zero of c_i. Find $\bar{\alpha} = \min(\alpha_i)$, and let I be the index for which $\bar{\alpha} = \alpha_I$. In other words, $\bar{\alpha}$ is the smallest positive first-order step to a constraint zero. If no constraint is decreasing along p, set the initial step $\alpha^{(0)}$ to α_u, and skip the remaining logic.
3. Estimate α_b, the step to the minimum of the barrier function along p, which satisfies:

$$(4.1) \qquad p^T\left(g(x + \alpha_b p) - \sum_{i=1}^{\ell} \frac{r}{c_i(x + \alpha_b p)} a_i(x + \alpha_b p)\right) = 0 ,$$

where $g(\cdot)$ is the gradient of F. The relationship (4.1) can be used to obtain a crude estimate of α_b, if two approximations are made. First, we assume that only the influence of the constraint $\bar{I}$ is significant in the location of α_b; this assumption is based on the idea that for small r, only the singularity along p closest to the starting point affects the local behavior of the barrier function. Second, it is assumed that the gradients of the objective function and $\bar{I}$-th constraint remain fixed locally. Under these conditions -- ignoring second-order terms, and all but the $\bar{I}$-th constraint -- (4.1) becomes:

$$g^T p - \frac{r\, a_{\bar{I}}^T p}{c_{\bar{I}} + \alpha_b\, a_{\bar{I}}^T p} \sim 0 ,$$

so that

$$\alpha_b \sim \frac{r}{g^T p} - \frac{c_{\bar{I}}}{a_{\bar{I}}^T p} .$$

Since

$$\bar{\alpha} = \frac{-c_{\bar{I}}}{a_{\bar{I}}^T p} ,$$

the estimate of α_b is given by:

$$\alpha_b \sim \frac{r}{g^T p} + \bar{\alpha} .$$

If $p^T g > 0$, i.e., p is not a descent direction for the objective function, then the assumption that only one constraint will have an effect, confined to the neighborhood of the singularity, is not justified, and we set $\alpha_b = \gamma\bar{\alpha}$, where $0 < \gamma < 1$.

4. If α_b exceeds α_u, let $\alpha^{(0)}$ be α_u; otherwise, $\alpha^{(0)} = \alpha_b$.

In this way, the initial step can be chosen based on the estimated decrease of the constraints if it seems that their effect will be significant in the location of the minimum of the barrier function.

4.2. Feasibility Check

For each step to be taken, the set of constraints is evaluated in order to assure that feasibility is never violated.

If any constraint, say the j-th, is non-positive at $x + \alpha^{(k)}p$, the secant step to the predicted zero of that constraint is computed, and the next estimate of the location of the minimum is computed as described above in Step (3), where the secant step $\bar{\alpha}_j$, given by:

$$\bar{\alpha}_j = \frac{-\alpha^{(k)}\, c_j(x)}{c_j(x + \alpha^{(k)}p) - c_j(x)} \quad ,$$

is used as $\bar{\alpha}$. This procedure is subject to the safeguarding requirement that the constraints not be evaluated at points that are too close together (see Gill and Murray, 1974; Brent, 1973, for a discussion of this aspect of safeguarded linear searches).

4.3. Normal Iteration of Modified Linear Search

The special functions are fitted during the iteration if a flag has been set to 'true'. The flag is set: (1) when the initial $\alpha^{(0)}$ was α_b, implying that the influence of some constraint is predicted to be significant in locating the minimum; and (2) when any negative constraint value is encountered during execution of the linear search, since it has then been demonstrated that the current iterate is in a region influenced by the singularities.

The special functions are fitted iteratively, using the same criteria for replacing points as the usual linear searches, until the particular convergence criteria are satisfied. There are a few subtle difficulties in that for small r, it may be difficult to locate the minimum because the distance from the singularity to the minimum may be less than the spacing required for constraint evaluations. However, a careful regulation of the tolerances involved, so that impossible accuracy is not sought, will assure that the process will work as desired.

4.4. Comparison with Usual Linear Searches

In order to determine whether the special linear searches are worthwhile, numerical experiments were carried out for several barrier functions, with varying values of r, the barrier parameter, and η, the linear search convergence parameter. For the cubic case where gradients are evaluated at every point, the linear search usually terminates when

$$|\bar{g}(x + \alpha p)| < \eta \, |\bar{g}(x)| ,$$

where $\bar{g}(\cdot)$ is the projected gradient of the function to be minimized; in the quadratic case, the linear search is usually terminated when the minimum of $\bar{F}$ is known to be bracketed in the interval $[a, b]$ and

$$\left| \frac{\bar{F}(x + \alpha p) - \bar{F}(x + ap)}{\alpha - a} \right| < \eta \, |\bar{g}(x)| ,$$

i.e., when the linearized approximation to the gradient at $x + \alpha p$ satisfies the same test as $\bar{g}(x + \alpha p)$ in the cubic case. There are other occasions when the normal linear search procedure will terminate, involving sufficient smallness of the interval of uncertainty, closeness to the maximum permitted step, etc.

Numerous runs (about 40) were made. For both the special functions and the usual polynomials, the same initial step was taken, and the same procedure was followed for determining the next point if a constraint became negative during the line search iteration. Hence, the only difference was in the use of the minima of special functions, rather than of cubic or quadratic polynomials, to yield the next point at which the function and constraints are to be evaluated as the linear search proceeds. In every case, use of a special function reduced the number of function and constraint evaluations; the reduction became progressively more significant as the value of η was reduced. The reduction in the number of function and constraint evaluations ranged from 7% to 20% when a Newton-type algorithm was used to carry out the unconstrained minimization, and from 12% to 24% for a quasi-Newton algorithm; thus, there was clear improvement with the approximation by special functions.

5. Conclusions

The extra work required in the linear search procedure to fit these special functions is small. Some of the housekeeping (checking for feasibility, etc.) must be carried out with barrier functions regardless of whether special functions are used or not. The formulations presented here allow calculation of the minimum of the fitted functions with the same information required to fit the usual polynomials. The singularity must be located through an iteration, but because of the special form of the iteration functions, we are able to obtain a highly accurate starting guess; in fact, in two of the three cases, the solution could be obtained from tables. The iteration functions are well-behaved, and Newton's method will usually converge to the desired accuracy within two iterations. Each iteration to locate the singularity requires evaluation of a transcendental function, but the subsequent reduction in the number of function and constraint evaluations required to locate a satisfactory approximation to the minimum of the barrier function seems ample justification for use of the special linear searches designed to minimize barrier functions.

References

1. R.P. Brent (1973), Algorithms for Minimization without Derivatives, Prentice-Hall, Englewood Cliffs, New Jersey.

2. A.V. Fiacco and G.P. McCormick (1968), Nonlinear Programming: Sequential Unconstrained Minimization Techniques, John Wiley and Sons, New York.

3. R. Fletcher (1974), Methods Related to Lagrangian Functions, in Numerical Methods for Constrained Optimization, P.E. Gill and W. Murray, eds., Academic Press, London and New York, pp. 203-213.

4. R. Fletcher and A.P. McCann (1969), Acceleration Techniques for Nonlinear Programming, in Optimization, R. Fletcher, ed., Academic Press, London and New York, pp. 203-213.

5. P.E. Gill and W. Murray (1974), Safeguarded Steplength Algorithms for Optimization Using Descent Methods Report NAC 37, National Physical Laboratory, Teddington, England.

6. L.S. Lasdon, R.L. Fox, and M.W. Ratner (1973), An Efficient One-Dimensional Search Procedure for Barrier Functions, Math. Prog., 4, pp. 279-295.

7. W. Murray (1969), Constrained Optimization, Report MA 79, National Physical Laboratory, Teddington, England.

8. W. Murray and M.H. Wright (1976), Trajectory Methods for Nonlinearly Constrained Optimization, to appear.

9. M. Osborne (1972), Topics in Optimization, Report 72-279, Computer Science Department, Stanford University, Stanford, California.

10. D.M. Ryan (1974), Penalty and Barrier Functions, in Numerical Methods for Constrained Optimization, P.E. Gill and W. Murray, eds., Academic Press, London and New York, pp. 175-190.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

14

1. REPORT NUMBER: SOL 76-18
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:

9

6

4. TITLE (and Subtitle): Efficient Linear Search Algorithms for the Logarithmic Barrier Function
5. TYPE OF REPORT & PERIOD COVERED: Technical Report
6. PERFORMING ORG. REPORT NUMBER:

10

7. AUTHOR(s): Walter Murray [redacted] Margaret H. Wright

15

8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0865 E(04-3)-326
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Operations Research, SOL, Stanford University, Stanford, CA 94305
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR-047-064

11

11. CONTROLLING OFFICE NAME AND ADDRESS: Operations Research Program Code 434, Office of Naval Research, Arlington, VA 22217
12. REPORT DATE: Sep[redacted]76
13. NUMBER OF PAGES: 50

12

55p.

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): This document has been approved for public release and sale; its distribution is unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number):
    - LINEAR SEARCH ALGORITHM
    - STEPLENGTH ALGORITHM
    - ONE-DIMENSIONAL MINIMIZATION
    - LOGARITHMIC BARRIER FUNCTION
    - CONSTRAINED OPTIMIZATION
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

    see next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

408765

JB

EFFICIENT LINEAR SEARCH ALGORITHMS FOR THE
LOGARITHMIC BARRIER FUNCTION

by
Walter Murray and Margaret H. Wright

Technical Report SOL 76-18

Linear search algorithms are developed for use when minimizing logarithmic barrier functions, whose one-dimensional behavior is in general modeled poorly by the low-order polynomial approximations of standard linear search procedures. The new methods are based on special approximating functions with a logarithmic singularity, and are designed to utilize the same information as procedures based on special approximating functions with a logarithmic singularity, and are designed to utilize the same information as procedures based on quadratic or cubic polynomials. Although the parameters of the special approximating functions depend nonlinearly on the available data, the determination of the parameters requires little additional work in comparison with polynomial fits. Use of the special approximating functions has led to a significant improvement in efficiency when minimizing logarithmic barrier functions, where efficiency is measured by the number of function (or function and gradient) evaulations required for termination of each linear search.